

# Acesso Multimodal em Dispositivos Móveis a Vídeo Georeferenciado através da Forma, Velocidade e Tempo

Sérgio Serra

Ana Jorge (\*)

Teresa Chambel

LaSIGE, Faculdade de Ciências, Universidade de Lisboa

(\*) and Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa  
Portugal

{sergioserra99,ananunesjorge}@gmail.com, tc@di.fc.ul.pt

---

## Resumo

*O vídeo está a tornar-se uma meio cada vez mais capturado, publicado e acedido na web, a partir de diferentes plataformas e dispositivos. Os utilizadores podem mais facilmente georeferenciar a informação que capturam e acedem, permitindo enriquecer a sua contextualização. A procura de vídeo tem sido limitada essencialmente a palavras-chave, ou a um conjunto de parâmetros com suporte muito limitado às dimensões espacial e temporal.*

*Propomos novos mecanismos para procurar e aceder a vídeos georeferenciados, onde as dimensões espacial e temporal são centrais, permitindo um acesso baseado na forma e velocidade das trajectórias inerentes a estes vídeos, usando uma interface multimodal num dispositivo móvel, baseada em gestos e movimento, e com o potencial de uma interacção mais natural, um maior envolvimento, sensação de presença e imersão. A avaliação com utilizadores, feita com um protótipo de alta fidelidade teve resultados positivos e encorajadores. Os utilizadores consideraram a maioria das funcionalidades úteis, satisfatórias, por vezes divertidas, e fáceis de usar. As diferentes opções e modalidades foram consideradas interessantes e adequadas a diferentes cenários de utilização que foram identificados e sugeridos. Também foram identificados alguns desafios para serem levados em conta em desenvolvimentos futuros, em direcção a um acesso interactivo mais flexível e eficaz, através de uma interacção mais natural com dispositivos móveis de per se, ou como segundos ecrãs em conjunto com ecrãs de maiores dimensões em TV ou ecrãs públicos.*

## Palavras-chave

*Vídeos Georeferenciados, Interacção Multimodal, Gestos e Movimento, Procura, Acesso e Navegação, Espaço e Tempo, Forma e Velocidade, Trajectórias, Dispositivos Móveis e Segundo Ecrã.*

---

## 1. INTRODUÇÃO

Todos os dias, uma grande quantidade de vídeos digitais é publicada na web. O vídeo está a tornar-se amplamente capturado, partilhado e acedido através de diferentes plataformas e dispositivos, que de forma crescente pode ser georeferenciado (em localizações pontuais e ao longo de trajectórias), permitindo enriquecer a sua contextualização. Apesar de muitos vídeos estarem disponíveis para ser procurados e acedidos, os mecanismos actuais e mais usados para os procurar e navegar são baseados num conjunto limitado de parâmetros como: palavras-chave, duração, e qualidade do vídeo, ignorando as suas dimensões temporais e espaciais. O vídeo tem um enorme potencial para imersão e os dispositivos móveis permitem aceder à informação enquanto ‘imersos’ na realidade, em qualquer lugar. Com a proliferação de dispositivos como: *smartphones*, *tablets* e mais recentemente *wearables*, poderíamos tirar partido dos sensores multimodais disponíveis para criar novas formas de encontrar e navegar vídeos georeferenciados através do espaço e do tempo (quer ao longo da sua duração e linha de tempo, quer o tempo da sua captura) e espaço, usando interfaces mais naturais, envolvendo gestos e a forma e a velocidade do

movimento, com o potencial de maior envolvimento, sensação de presença e imersão, ao aceder aos vídeos.

Este trabalho surge no seguimento de um outro trabalho prévio no contexto do Sight Surfers [Noronha12] [Ramalho13a,13b], uma aplicação web interactiva para a partilha, visualização e navegação de vídeos 360° georeferenciados e interactivos, como hipervídeos, incluindo *tours/circuitos* em cidades ou actividades mais radicais, como corridas de kart. Estes vídeos podem ser experienciados de forma mais imersiva, isoladamente, ou sincronizados com um mapa ao longo do percurso. Para permitir um suporte mais completo à dimensão espaciotemporal em vídeos georeferenciados, e com o objectivo de aumentar a imersão associada a uma experiência sensorial aumentada, pretendemos criar mecanismos mais ricos para a procura, visualização e navegação interactivas através de modalidades de interacção mais naturais.

Neste artigo, descrevemos o nosso trabalho desenvolvido nesta direcção. A próxima secção evidencia os principais desafios e oportunidades, e descreve o trabalho relacionado mais relevante. Em seguida, apresenta-se o modelo conceptual e as opções de design para o acesso multi-

modal a vídeo georeferenciado no espaço e no tempo, e dispositivos móveis, que são demonstrados em protótipos, e avaliados na secção seguinte. Foi realizada uma avaliação com utilizadores e protótipos de alta fidelidade, para aferir a usabilidade e o nível de aceitação, focando a utilidade, a satisfação e a facilidade de uso. Finalmente, o artigo termina com conclusões e perspectivas para trabalho futuro.

## 2. TRABALHO RELACIONADO

Os desafios para este trabalho incluem a concepção e a realização de uma interface interactiva adequada, capaz de capturar e expressar as dimensões temporal e espacial, e permitir a representação da velocidade e da forma das trajectórias ao longo do tempo, através de um modo intuitivo, simples, eficaz e natural de procurar, apresentar e navegar os vídeos em ambientes móveis. É simultaneamente um desafio e uma oportunidade, porque os utilizadores não estão habituados a procurar e a navegar nestas dimensões. Mas a tecnologia já permite capturar o movimento em dispositivos móveis de formas que têm o potencial de suportar modos de interacção mais naturais, envolvendo o tempo, a forma e a velocidade, permitindo prosseguir em direcção a experiências mais imersivas.

### 2.1 Georeferenciação em Bibliotecas de Vídeo

A maioria das bibliotecas de vídeo e websites como o YouTube ou o Vimeo baseiam-se em palavras-chave e têm quando muito um suporte muito limitado ao acesso a vídeo com base nas dimensões espacial e temporal. Rego et al. [Rego07] desenvolveram a VideoLIB, uma biblioteca digital que melhora a recuperação de vídeo através de operadores espaciais e temporais, baseados nas normas de metadados Dublin Core e MPEG-7. Os critérios de pesquisa incluem: acção (o quê), pessoa (quem), tempo (quando) e local (onde); e usa operadores como: antes, durante e depois para definir intervalos de tempo. Isto permite efectuar pesquisas como: “procura os clips de vídeo da Madonna que foram produzidos fora dos estados unidos da américa durante os anos 1990s”. Utiliza um formulário numa interface baseada em texto, sem o recurso a mapas, e os vídeos são considerados como um todo, e as trajectórias e a velocidade não são tidas em consideração.

### 2.2 Acesso a Fotos e Vídeos através de Mapas

Há algumas abordagens à procura e navegação de vídeos, e especialmente de fotos, que utilizam mapas. O Google Street View é um visualizador de fotos 360° que usa uma projecção esférica da imagem e geolocalização, mas não suporta vídeo, nem vistas alternativas, inclusivé criadas/capturadas pelos utilizadores, dos diferentes espaços. Panoramio (.com) é um website de partilha de fotos georeferenciadas acedido como um *layer* do Google Earth e Google Maps. Os utilizadores podem efectuar pesquisas baseadas em texto ou navegar nos mapas, e ver as fotos tiradas por outros utilizadores, com base na localização. As fotos são apresentadas juntamente com um mapa que evidencia a sua localização, quer como uma colecção resultante de uma pesquisa, quer de forma individualizada: uma a uma. Existem filtros para evidenciar as fotos mais populares, as mais recentes, locais famosos,

em interiores, que através de um separador próprio para as fotos filtradas quer aumentando a dimensão destas fotos de entre todas as que são apresentadas no mapa.

Finsterwald et al. [Finsterwald 12] desenvolveram o MOMA (The Movie Mashup Application), como um serviço public web baseado em mapas para pesquisar filmes baseados na localização, combinando recursos de georeferenciação e processamento de texto, integrando informação oriunda da DBpedia, GeoNames e Wikipedia. Através da sua interface gráfica, permite procurar e navegar uma base de filmes por realizador, localização por texto, através de áreas poligonais no mapa, a partir de localizações extraídas de títulos de filmes, comparar distribuição das pesquisas e, utilizando uma versão móvel, permite pesquisar filmes em que acção tem lugar numa zona perto da localização actual do utilizador. Apesar de os mapas constituírem uma forma natural de representar informação georeferenciada, e de os vídeos frequentemente envolverem trajectórias, a maioria das soluções existentes só permitem que os utilizadores publiquem e acedam a vídeos com base numa localização GPS única (geralmente a posição inicial). Hao et al. [Hao11] apresentam vídeos gerados pelos utilizadores relacionados com áreas geográficas numa interface baseada num mapa. Eles focam-se na selecção automática de *keyframes* para representar os vídeos, e na determinação da sua localização, para as colocar no mapa. Desta forma, enfatizam os pontos de interesse (exo. um monumento) que são capturados nos vídeos, e não tanto as trajectórias.

### 2.3 Interacção Háptica para acesso a *Media* em Dispositivos Móveis

Em termos de pesquisa interactiva espacial e háptica em ambientes móveis, nos últimos anos, notou-se uma crescente popularidade das interfaces gestuais e das aplicações em segundo ecrã. Lei & Coulton [Lei09] implementou uma aplicação controlada por gestos que age como uma varinha, usando sensores móveis. Permite pesquisa de pontos de interesse (POI) por proximidade ou remotas, baseada na orientação da varinha, como um ‘Flashlight’ interactivo e espacial, e a possibilidade dos utilizadores criarem conteúdo adicional para um POI particular como fotos etiquetadas com a localização do POI e a direcção de onde a foto foi tirada. As fotos podem ser filtradas com base num ângulo de visão pretendido num contexto do mundo real.

Premraj et al. [Premraj10] apresentaram o iWalk, uma ferramenta que permite exploração multimédia de dados georeferenciados através: do movimento, para movimentar no espaço digital de uma colecção; e gesto, para a manipulação directa dos dados (exo. *select, go to next, zoom*). Eles fizeram experiências com colecções de fotos e sons georeferenciados, e com uma colecção de museu não georeferenciada, onde o utilizador definia o mapeamento entre os espaços digital e físico. A sua abordagem faz uso de algoritmos de visão por computador executados em tempo-real em câmaras comerciais comuns.

## 2.4 Interação e Acesso a Vídeo com Segundo Ecrã

Os dispositivos móveis podem também actuar como segundos ecrãs [Courtois12] para complementar e interagir com ecrãs de dimensões superiores como TVs ou até ecrãs em locais públicos. MobiToss, é uma aplicação criada por [Scheible08], que permite a criação e partilha de arte multimédia num grande ecrã público, através de dispositivos móveis com acelerómetros. Os utilizadores podem tirar uma foto ou capturar um vídeo no dispositivo móvel, e “atirá-los” para um grande ecrã público, com um gesto, para visualização e manipulação, através de *tilting*. Os clips criados pelos utilizadores são aumentados ou enriquecidos pelo sistema com items como música ou nomes de marcas e enviados de volta para os seus *smartphones*, como artefactos pessoais do evento.

A avaliação preliminar com utilizadores mostrou que a captura e o atirar do conteúdo móvel para o grande ecrã, bem como a sua manipulação através de gestos na criação da peça de arte foi percebida como uma actividade intuitiva e divertida. Eles gostaram e envolveram-se na experiência, e apreciaram levar algo com eles, especialmente algo com um cariz artístico. Mas a aplicação necessita de melhoramentos, aplicando um conjunto de efeitos de vídeo mais equilibrado, adicionando interacção em grupo e uma interface mais intuitiva, para acomodar

os diferentes gestos que os utilizadores usam para atirar, e aumentar a percepção sobre o que está a acontecer em cada momento. Este trabalho explora gestos naturais com dispositivos móveis para manipular fotos e vídeos no segundo ecrã, mas não explora as dimensões espacial e temporal dos vídeos (e das fotos). E nenhum dos trabalhos relacionados encontrados abordam velocidade e trajectórias em vídeo, como nos propomos fazer.

## 3. ACESSO A VÍDEO ATRAVÉS DO ESPAÇO, VELOCIDADE E TEMPO

As dimensões espacial e temporal são aqui tidas em conta essencialmente na localização dos vídeos, na forma e velocidade das trajectórias filmadas, e no tempo dentro (exo. ao longo da sua duração) e fora do vídeo (exo. quando foi filmado). Para explorar a interacção e a experiência dos utilizadores com estas dimensões em vídeos georeferenciados, foi desenvolvido um protótipo de alta fidelidade para Android (Fig.1 e 2). Foram concebidas variantes para diferentes cenários de utilização, explorando o uso de sensores e sistemas de localização, na sequência de trabalho preliminar prévio com protótipos de baixa fidelidade [Serra14]. Em seguida, apresentamos o racional subjacente às principais opções de design.



Figura 1. Diferentes modalidades de entrada, através de: a) Gestos (velocidade); b) Toque (velocidade); c) Em “viagem” ou deslocação, exo. num carro; d) Gesto (forma); e) Toque para procura por forma num mapa; e f) e no tempo, por vídeos mais antigos.

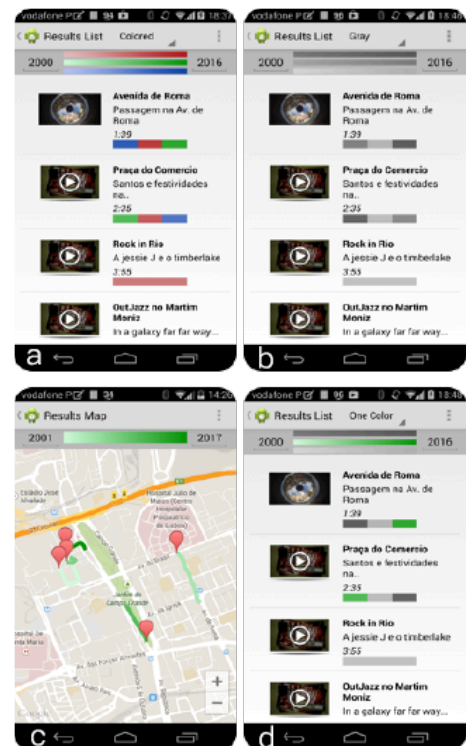


Figura 2. Diferentes modalidades de saída: a) Lista com linha de tempo de 3 cores; b) Lista com linha de tempo de tons de cinzento; c) Resultados num Mapa representando vídeos através das suas trajectórias; e d) Lista com linha de tempo de 1 cor.

### 3.1 Através de Toque - com o dedo

Esta é a interface mais convencional, que permite desenhar as formas de pesquisa no ecrã através de toque (Fig.1d), ou inclusive num *touch pad* ou rato num portátil ou num PC. A velocidade de realização deste desenho pode também ser capturada para pesquisar apenas por velocidade, ou por ambas: forma e velocidade. Este tipo de interacção pode ser mais familiar e permitir uma maior precisão do que as seguintes, especialmente se o utilizador tiver uma mão livre para fazer esta interacção.

### 3.2 Através do Gestos - com o telemóvel

Esta modalidade pode se usada movendo o telemóvel (*smartphone*) através de um gesto, para desenhar a forma (Fig.1d) ou demonstrar a velocidade (Fig.1b). A velocidade do gesto é capturada usando o acelerómetro, mas este sensor não permitiu capturar a forma do movimento com precisão suficiente, do modo previsto no protótipo de baixa fidelidade. Assim sendo, foi também usado o giroscópio para capturar a forma do movimento, e o movimento passou a ser efectuado através da inclinação do telemóvel para desenhar a forma como se o cursor fosse uma pequena bola a deslocar-se em cima do ecrã, de acordo com a inclinação feita (Fig.1d).

Esta interacção pode ser efectuada com a mão que segura no telemóvel, deixando a outra mão livre, e tem o potencial para uma modalidade mais natural ou imersiva para seleccionar os vídeos pretendidos, a serem visualizados no telemóvel ou num ecrã de maiores dimensões como uma TV. Neste contexto, os espectadores estão habituados a interagir com um comando numa só mão, enquanto se mantêm focados no vídeo no grande ecrã.

### 3.3 Através da Deslocação - em movimento

Quando em deslocação, os utilizadores frequentemente viajam de carro, comboio, metro, avião, etc., ou andam e correm. Em alternativa ou para além do uso da localização actual para aceder a vídeos capturados na mesma localização em que me encontro, pode ser interessante, especialmente quando não se vai a guiar, aproveitar a oportunidade de ver vídeos que foram capturados a uma velocidade análoga à que nos estamos a deslocar, permitindo experienciar o visionamento de forma mais imersiva, fazendo coincidir a velocidade do que estamos a ver com aquela que estamos a sentir fisicamente na realidade (Fig.1c). Isto pode ser particularmente impactante em vídeos de alta velocidade, por exemplo em actividades mais radicais, como corridas de carros, motocross ou sky. Mas também permite por exemplo procurar por vídeos capturados na mesma trajectória que estou a fazer a pé, a uma velocidade análoga, mas há uns anos atrás, o que me permitirá ir comparando a minha experiência actual com o que se podia experienciar nessa altura na mesma localização. Como nos casos anteriores, quer a velocidade quer a forma da trajectória podem ser capturados nesta modalidade. A velocidade e a localização poderão ter um maior potencial de imersão na visualização imediata de vídeos, para obter vídeos análogos ou relacionados com a minha realidade no momento. Mas capturar a trajectória pode também ser interessante para procurar por vídeos que tenha trajectórias semelhantes, mesmo que filmados

em localizações diferentes, por exemplo outras corridas de carros com trajectos semelhantes em qualquer parte do mundo. Enquanto a captura de gestos se baseia em sensores, esta funcionalidade de deslocação, ou viagem, baseia-se em serviços de localização como o GPS.

### 3.4 Onde - no Mapa or na Localização Actual

Uma vez que os vídeos na nossa aplicação Sight Surfers são georeferenciados, as pesquisas podem ser feitas com base na localização. Os utilizadores podem usar um mapa para seleccionar as localizações (Fig.1c), ou podem capturar a localização actual, para as pesquisas (para além de poderem especificar os locais, como por exemplo cidades, através de uma interface mais tradicional baseada em texto).

### 3.5 Em qualquer Lugar

Os vídeos podem também ser procurados independentemente da sua localização. Para tal, não se faz recurso a mapas, e só a velocidade e a forma são desenhadas no ecrã ou no ar, ou capturadas em deslocação, ou viagem, sem georeferenciação.

### 3.6 Resultados em Mapas e Listas

Os vídeos resultantes podem ser apresentados como trajectórias num mapa (Fig.2c), onde cada trajectória pode ser vista como uma linha de tempo (*timeline*) do vídeo sincronizada com o vídeo a ser visionado, como no Sight Surfers [Noronha12]. É esta a opção por omissão quando a pesquisa é baseada na localização. Mas os resultados também podem ser apresentados de forma independente da sua localização, por exemplo numa lista onde a velocidade (Fig. 2abd) e/ou a forma podem ser evidenciadas na *timeline* de cada vídeo. Note ainda que os utilizadores podem alternar entre as vistas de mapa e lista e seleccionar o que pretendem ver, sobre os mesmos resultados. Os resultados das pesquisas são apresentados com diferentes alternativas de design, evidenciando diferentes aspectos e informações sobre os conteúdos retornados, em termos de forma, velocidade e tempo, quer em mapa ou em lista.

#### 3.6.1 Percepção de velocidade

A velocidade pode variar ao longo de um vídeo, pelo que se devem apresentar primeiro, como resultado, os vídeos que têm a velocidade procurada (com uma tolerância) por mais tempo, e deve dar-se a percepção de quando no vídeo a velocidade é como procurado, e quando é maior e menor. A Fig.2abd apresenta três alternativas de design para a apresentação da velocidade nas linhas de tempo dos vídeos, através de: a) 3 Cores: verde para a velocidade procurada, vermelho para mais rápido e azul para mais lento (RGB); b) Tons de Cinzento: um tom intermédia para a velocidade procurada, mais escuro para mais rápido e mais claro para mais lento; and c) 1 Cor: verde para a velocidade procurada e dois tons de cinzento para mais rápido e mais lento, como em b), permitindo um maior contraste e destaque para a cor procurada.

#### 3.6.2 Percepção de forma

A forma é apresentada por omissão na vista de mapa, mas também pode ser apresentada na lista, como na Fig.2c, onde as linhas de tempo de cada vídeo tomam a

forma da trajectória correspondente. A percepção de velocidade é opcional em todas as linhas de tempo: mapa e lista, com ou sem forma.

### 3.6.3 Percepção e procura por tempo

Para além de se apresentar a duração junto de cada vídeo, e o tempo corrente na linha de tempo, o tempo em que o vídeo foi capturado é representado pela cor, quer da linha de tempo nas listas (Fig.2abd) quer das trajectórias nos mapas (Fig.2c), para representar a sua idade. Optamos por perder saturação para os vídeos mais velhos, e manter a vivacidade da cor para os mais novos, de forma proporcional.

No topo do ecrã, uma linha de tempo apresenta a escala de tempo de referência, colorida de forma correspondente à “idade” dos vídeos. A menor saturação à esquerda, para os mais antigos, que vai gradualmente aumentando para a direita, em direcção aos mais recentes. As etiquetas de tempo (um ano no exemplo da Fig.2) que representam o intervalo de tempo são apresentadas em caixas de texto à esquerda e direita da linha de tempo. Esta linha torna explícito o mapeamento entre as cor e o tempo, permitindo a identificação dos vídeos mais recentes e mais antigos no intervalo de tempo em causa. Esta linha de tempo pode ter apenas uma cor, quando a velocidade não é apresentada (por exemplo, apenas verde no mapa da Fig.2c), ou ter três barras, uma por cada categoria de velocidade (por exemplo, tons de cinzento ou RGB nas listas das Figs.2ba) para evidenciar como as diferentes cores “envelhecem”.

Nesta vista, os utilizadores podem procurar ou filtrar vídeos por tempo: quer seleccionando por toque as duas caixas de texto das etiquetas, e escrevendo o ano pretendido num teclado virtual; quer usando gestos para incrementar ou decrementar o seu valor, desenhando uma espiral com o dedo no ecrã (o gesto evidenciado no desenho na Fig1.d). Uma espiral no sentido dos ponteiros dos relógios significa avançar no tempo, enquanto que no sentido contrário significa recuar. Este mapeamento foi escolhido com base numa hipótese de que seria o mapeamento mais natural, considerando a metáfora dos relógios de ponteiros.

## 4. AVALIAÇÃO COM UTILIZADORES

Foi realizada uma avaliação com utilizadores para avaliar as funcionalidades concebidas em termos da sua utilidade, experiência de utilização e usabilidade, para investigar sobre as alternativas mais eficazes e preferidas, e sobre a sua aplicação em cenários de uso reais.

### 4.1 Método

Foi realizada uma avaliação baseada em tarefas, executadas pelos utilizadores participantes, recorrendo a Observação e Entrevistas semi-estruturadas, depois de se explicar o objectivo da avaliação, o conceito subjacente ao contexto da aplicação Sight Surfers, e as funcionalidades em avaliação no protótipo de alta fidelidade. A ordem das variantes foi alterada para cada utilizador, para que esta não viesasse os resultados, seguindo uma amostragem por medidas repetidas, onde todos os utilizadores testaram todas as variantes do protótipo.

No final de cada tarefa, os utilizadores atribuíram uma classificação baseada no questionário USE (Usefulness, Satisfaction, and Ease of use [Lund01], que captura as dimensões essenciais de utilidade, experiência de utilização e usabilidade) utilizando uma escala de 1-5, para cada funcionalidade interactiva testada, em todas as suas variantes, e encorajados a tecer comentários e apresentar sugestões. As avaliações tiveram lugar no campus da nossa universidade, maioritariamente em espaços interiores, onde os utilizadores, acompanhados pelo avaliador, podiam sentar-se ou caminhar.

### 4.2 Participantes

A avaliação contou com 9 utilizadores participantes com idades compreendidas entre 18 e 27 anos (22,5 em média, 3 do sexo feminino e 6 do sexo masculino). Todos os utilizadores haviam no mínimo terminado a escola secundária, 3 estavam na área de informática, os restantes de áreas diversas, todos tinham um *smartphone* usado diariamente para aceder a informação, e 9 faziam diariamente pesquisas e visionamentos de vídeos mas maioritariamente em PCs, por vezes em *tablets*, e mais raramente em telemóvel.

### 4.3 Resultados

Os principais resultados são sumarizados através dos valores médios do USE e dos comentários mais significativos, em cada categoria de funcionalidades.

#### 4.3.1 Procura por velocidade

Foi pedido aos utilizadores participantes que “procurassem vídeos por velocidade, usando toque e gestos”. Através de toque: foi considerado bastante divertido, muito fácil de usar, e útil pela maioria dos utilizadores (U: 3.67; S: 3.78; E: 4.44). Exemplo de comentário: “*A pesquisa georeferenciada é sem dúvida útil em muitas situações, por exemplo para conhecer certas zona de uma cidade*”, “*Eu imagino um atleta a usar esta funcionalidade para procurar pistas de corrida*”. Através de gestos: teve classificações ligeiramente mais baixas, especialmente na facilidade de uso (U: 3.22; S: 3.22; E: 3.67), nesta nova modalidade, menos familiar.

Finalmente, foi pedido aos utilizadores que “procurassem vídeo usando a velocidade de ‘viagem’ ou deslocação”. Foi-lhes pedido que “procurassem por vídeos com uma velocidade correspondente a um passo rápido”, que eles realizaram andando. Os resultados foram satisfatórios (U: 3.43; S: 3.0; E: 4.11), mas alguns utilizadores mencionaram que esta funcionalidade poderia ser mais útil quando em viagem por exemplo de carro ou comboio, e que se tinha sentido pouco à-vontade a caminhar rapidamente para a avaliação, não se sentindo suficientemente confortáveis e relaxados para depois ver os vídeos, enquanto se mantinham a caminhar com a mesma velocidade. Outros ambientes e contextos de deslocação e viagem serão testados em avaliações futuras.

#### 4.3.2 Procura por forma

Foi solicitado aos utilizadores que “procurassem por uma dada forma numa localização específica no mapa através de toque (usando o dedo)”. Esta funcionalidade foi bem recebida (U: 3.44; S:3.11; E:3.44): os participantes apre-

ciaram ter a capacidade de georreferenciar os trajectos representados pelas formas, e encontraram vários cenários onde poderia ser usada: por exemplo: “*Como atleta, eu poderia usar esta pesquisa para explorar melhor pistas de corrida com certos formatos, e vê-las no mapa*”, “*A procura georreferenciada é sem dúvida útil em muitas situações, por exemplo para conhecer melhor certas partes da cidade*”.

A tarefa “Procure por vídeo com o formato em ‘u’, usando o telemóvel para desenhar a forma com um gesto” foi considerada muito difícil e pouco útil, tendo sido a que teve pior aceitação por parte dos participantes (U: 1.89; S: 2.22; E: 2.22), “*Eu não considero que seja prático de realizar esta procura enquanto estou a fazer outras actividades, como a andar ou até a falar com outra pessoa*”. O equilíbrio entre a velocidade e a sensação de controlo é na nossa opinião longe de ser óptima, no estado actual do protótipo, não nos permitindo criar o tipo de experiência que tínhamos em mente para esta funcionalidade. Para além disso, os utilizadores não estão habituados a usar o telemóvel para este tipo de acção ou interacção, tendendo a preferir usar o dedo para desenhar.

#### 4.3.3 Resultados em Mapas e Listas

Em geral, os utilizadores preferem a forma mais familiar de apresentar vídeos numa **lista** (U: 4.0; S: 3.56; E: 4.44), por ser fácil e simples, “*É a forma mais rápida de ver vídeos, se a localização geográfica não for relevante*”. De qualquer forma, o **mapa** também teve boa aceitação (U: 3.44; S: 3.11; E: 3.78), e a maioria dos participantes referiram que estes eram úteis quando se queria procurar por trajectórias georeferenciadas, ou por vídeos em locais específicos, ou quando nos queremos aperceber das localizações e dimensões das trajectórias subjacentes aos vídeos. Também comentaram que: “*É interessante ver as trajectórias dos vídeos [no mapa] sem a necessidade de visionary o vídeo*”.

As principais preocupações mencionadas dizem respeito à percepção da quantidade de vídeos recuperados e a sua representação no caso de se uma quantidade muito elevada. A filtragem dos resultados não estava no âmbito deste teste, mas está alinhada com as nossas próprias questões de investigação mais abrangentes, e já abordadas inclusivamente em trabalhos anteriores [Ramalho13a,13b].

Houve ainda sugestões no sentido de ter uma mistura das vistas de lista e mapa, onde ao navegar uma lista de vídeos, poderíamos ver a sua localização num mapa; e quando passássemos o cursor por cima de uma trajectória no mapa, poderíamos ver informação mais detalhada sobre os vídeos nessa trajectória numa lista (com velocidade, duração, uma imagem representativa do vídeo, título, data etc.).

Em termos de **percepção e procura por tempo**, as linhas de tempo (*timelines*) foram consideradas úteis, satisfatórias, e fáceis de usar para encontrar a velocidade procurada nos vídeos resultantes. Apresentamos três designs alternativos, e a variante preferida foi a de “1 cor” (tabela 1). A maioria dos utilizadores preferiram este modo (U: 4.11; S: 4.11; E: 4.44) em que se evidencia a velocidade

procurada numa cor (nesta caso, em verde), enquanto as outras velocidades são menos evidenciadas (em cinzento), por ser mais fácil de usar. Em seguida preferiram a variante de “3 cores” (U: 3.44; S: 3.44; E: 3.6), que consideraram útil e bastante satisfatória, mas mais difícil de usar: “*A linha de tempo com uma cor é uma forma fácil e directa de encontrar a velocidade procurada. A linha de tempo com 3 cores é mais difícil de perceber e torna a interface menos clara, mais cheia e berrante*”. A variante em tons de cinzento foi considerada mais difícil de usar e menos útil, satisfatória e até menos divertida (U: 2.56; S: 2.44; E: 2.67). Estas preferências em relação aos designs alternativos para a percepção de velocidade condizem com as nossas expectativas, que correspondiam às nossas hipóteses descritas na fase de design.

designs	Utilidade		Satisfação		Facil. d/ Uso	
	M	$\sigma$	M	$\sigma$	M	$\sigma$
1 Cor	4,11	0,31	4,11	0,5	4,44	0,83
Cinzas	2,56	0,68	2,44	0,5	2,67	0,67
3 Cores	3,44	0,76	3,44	0,68	3,44	0,88

**Table 1: Percepção de velocidade em trajectórias nas linhas de tempo.**

Depois da procura, foi solicitado aos utilizadores que filtrassem os resultados “recuando no tempo para vídeo mais antigos”, em duas modalidades de interacção: usando o teclado virtual para escrever o ano, e gestos através de toque. Os resultados indicam que os utilizadores gostaram de ambas as modalidades para filtrar os resultados pelo tempo, no entanto, consideraram o teclado ligeiramente mais útil e fácil de usar (U: 3.67; S: 3.22; E: 4.0), enquanto que os gestos foram considerados mais satisfatórios e divertidos (U: 3.33; S: 3.67; E: 3.44). Numa perspectiva mais geral, os utilizadores consideraram que as duas modalidades se complementam, sendo os gestos mais rápidos e o teclado mais preciso.

## 5. CONCLUSÕES E TRABALHO FUTURO

Neste artigo apresentou-se a motivação e as opções de design para o acesso multimodal e móvel a vídeo georeferenciado, no espaço e no tempo, num contexto maioritariamente de conteúdo gerado pelos utilizadores. Estas ideias foram materializadas num protótipo de alta fidelidade que foi desenvolvido para Android e que foi avaliado com utilizadores, em termos da sua utilidade, facilidade de uso, satisfação e preferências.

O nosso foco assentou em três dimensões principais: forma, velocidade e tempo, e explorou a multimodalidade potenciada pelos telemóveis mais recentes (*smartphones*) para conceber novas formas de aceder e navegar em vídeos georeferenciados, com base em toque, gestos e movimento.

A avaliação com utilizadores foi encorajadora. Os utilizadores consideraram a maioria das funcionalidades bastante satisfatórias, até divertidas, e fáceis de usar, e as diferentes opções de design e modalidades de interacção foram consideradas interessantes e adequadas para

diferentes cenários de utilização que foram identificados. A utilidade e a facilidade de uso foram mais rapidamente associadas às modalidades mais familiares, enquanto que os gestos e o movimento, ou deslocação, potencialmente mais naturais foram considerados bastante satisfatórias e por vezes até divertidas. Em particular, os gestos para navegar no tempo, através de vídeos mais antigos e mais recentes, foram muito bem recebidos pelos participantes, e foram apontados como uma forma eficaz e rápida de procurar e navegar no tempo, que podia combinar bem com uma modalidade mais precisa, versátil e tradicional de introduzir um tempo específico, usando um teclado.

Por outro lado, a funcionalidade menos apreciada foi a procura por forma através de gestos com o telemóvel, que por razões técnicas assentou numa implementação que não permitia tanto controlo e precisão como havia sido concebido em baixa fidelidade. Adoptou a inclinação do telemóvel, em vez de permitir executar um gesto no ar com o telemóvel na mão, desenhando no ar a forma pretendida, o que também permitiria uma melhor captura da velocidade do movimento, para a procura também por velocidade. Esta foi considerada uma limitação do actual protótipo logo à partida, e ela veio a reflectir-se nos resultados, de acordo com a nossa expectativa.

Outro aspecto que pode vir a ser melhorado refere-se às condições de avaliação em cenários mais realistas, em que os utilizadores possam sentir mais de perto a necessidade ou os benefícios de usar trajectórias e a velocidade, para aceder aos vídeos nas diferentes modalidades e na “vida real”, como as que alguns identificaram nos comentários, mas que a maioria não está habituada a vivenciar.

Os próximos passos incluem o refinar e a extensão de algumas das soluções actuais, baseados no retorno obtido dos utilizadores, e explorar novas formas de navegar, por exemplo, no tempo, e também no próprio visionamento imersivo do vídeo. A área dos segundos ecrãs é também uma área que estamos a abordar, e algum trabalho está já a ser desenvolvido para se poder, por exemplo, “atirar” o vídeo para um ecrã de maiores dimensões, como uma TV. No grande ecrã pode-se visionar o vídeo de uma forma mais imersiva, enquanto se pode continuar a usar o dispositivo móvel para ajudar a orientação, a percepção e a navegação, inclusivé com a sincronização com um mapa e as opções de selecção, e para aceder a conteúdo mais detalhado ou relacionado. Estas direcções de investigação alinham com os desenvolvimentos mais recentes e as tendências na área de TV interactiva e vídeo online, crescentemente acedido através de um ecossistema com vários dispositivos, incluindo os móveis [Obrist15].

## 6. AGRADECIMENTOS

Este trabalho foi parcialmente suportado pela FCT através do Projeto Estratégico LaSIGE, com referência PEst-OE/EEI/UI0408/2014; o Projecto de investigação ImTV com referência UTA-Est/MAI/0010/2009; e a bolsa de doutoramento FCT dentro do programa UT Austin|Portugal com referência SFRH/BD/51798/2011.

## 7. REFERÊNCIAS

- [Courtois12] Courtois, C., D’heer, E. Second screen applications and tablet users: constellation, awareness, experience, and interest. *Proc. EuroITV’12*, 2012, ACM Press, 153-156.
- [Finsterwald12] Finsterwald, J. M., Grefenstette, G., Law-To, J., Bouchard, H., and Mezaour, A.D. The Movie Mashup Application MoMa: Geolocating and Finding Movies. *Proc. GeoMM’12*, 2012 ACM, 15-18.
- [Hao11] Hao, J., Wang, G., Seo, B., and Zimmermann, R. Keyframe Presentation for Browsing of User-generated Videos on Map Interfaces. *Proc. ACM MM’11*, 2011, ACM Press, 1013-1016.
- [Lei09] Lei, Z., and Coulton, P. A. Mobile Geo-wand Enabling Gesture Based POI Search an User Generated Directional POI Photography. *Proc. ACE’09, Int. Conf. on Advances in Computer Entertainment Technology*, 2009, ACM, 392-395.
- [Lund01] Lund, A.M. Measuring usability with the USE questionnaire. *Usability and User Experience*, 8(2), 2001.
- [Norman93] Norman, D. *Things that Make us Smart*. Addison Wesley Publishing Company, 1993.
- [Noronha12] G. Noronha, C. Álvares, T. Chambel. Sight Surfers: 360° Videos and Maps Navigation. *Proc. Geo MM’12*, 2012, ACM Press, 19-22.
- [Obrist15] Obrist, M., Cesar, P., Geerts, D., Bartindale, T., and Churchill, E. Online Video and Interactive TV Experiences. *Interactions*, ACM, Sep-Oct 2015, pp. 33-37.
- [Premraj10] Premraj, V., Schedel, M., and Berg, T.L. iWalk, A Tool for Interacting with Geo-Located Data Through Movement and Gesture. *Proc. ACM MM’10*, 2010, ACM Press, 1059-1062.
- [Ramalho13a] Ramalho, J. and Chambel, T. Windy Sight Surfers: sensing and awareness of 360° immersive videos on the move. In *Proc. of EuroITV 2013*, Como, Italy. 107-116.
- [Ramalho13b] J. Ramalho, Chambel, T. Immersive 360° Mobile Video with an Emotional Perspective. In *Proc. ImmersiveMe’2013* (ACM Press 2013) 35-40.
- [Rego07] Rego, A., Baptista, C., Silva, E., Schiel, U., and Figueirêdo, H. VideoLib: a Video Digital Library with Support to Spatial and Temporal Dimensions. In *Proc. SAC’07*, 2007, ACM, 1074-1078.
- [Scheible 08] Scheible, J., Ojala, T., and Coulton, P. Mo-biToss: A novel gesture based interface for creating and sharing mobile multimedia art on large public displays. In *Proc. ACM MM’08*, 2008, ACM Press, 957-960.
- [Serra14] Serra, S., and Chambel, T. Towards Multimodal Consumption of Georeferenced Mobile Video Using Shape and Speed. In *Proc. of 2nd Int. Workshop on Interactive Content Consumption* at ACM TVX 2014.